

発声障がい者のための 分析合成を用いたモーラ単位音声出力システムの開発

佐藤 春樹† 岩野 公司†
東京都市大学†

1. はじめに

病気などで発声が困難になった人（発声障がい者）のコミュニケーションを支援するため、発声の代替手段として、ポータブルな PC 端末を用いた音声出力装置「VOCA（Voice Output Communication Aid）」の開発が進められている。我々はこれまでに、モーラ（かな一文字分に相当する音声単位）を単位として音声出力を行う方式の VOCA の開発を行っている[1]。この VOCA には、(1)モーラごとに音声を出力できるため、相手の発話に対して即座に応答できる、(2)モーラごとに音声の韻律（抑揚や継続時間）を操作できるため、様々な韻律表現ができる、といった利点を備えている。しかし、従来 VOCA では、モーラごとの抑揚の変化のパターンが限定されていることや、出力音声の個性が固定されていることなど、表現力に関する問題を抱えていた。

そこで本研究では、これらの問題の解決を図った、リアルタイムの分析合成器を利用した VOCA の提案を行い、その性能評価を行う。

2. 従来 VOCA の特徴と課題

従来 VOCA[1] の入力画面の様子を図 1 の左に示す。画面左側に子音のパネル、右側に母音のパネルが配置されており、これらを順に押すことで 1 モーラの音声を出力することができる。その際、母音のパネルを特定の方向にドラッグすることで抑揚を操作することができる（図 1 右）。表現できる抑揚のパターンは、「平坦」「平坦で低い」「平坦で高い」「上がり調子」「下がり調子」の 5 種類である。また、母音パネルを押している間は音声が出力され続けるので、その時間でモーラ音声の継続時間を調整できる。また、この VOCA では別の音声合成システムで作成された男性話者 1 名による音声が入力源として登録されており、その話者を変更する機能は有していない。この VOCA の表現力を向

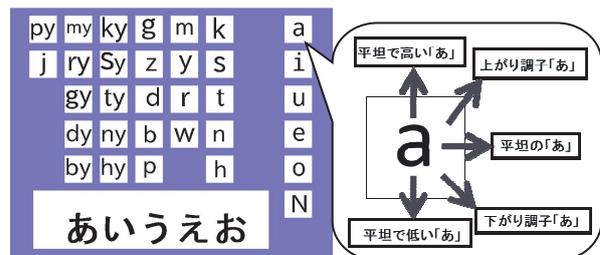


図 1 従来 VOCA[1]の入力画面（左）と母音パネルによる抑揚の操作方法（右）

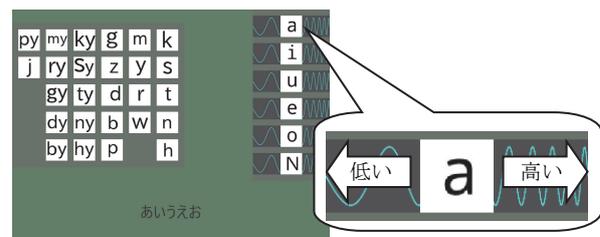


図 2 提案 VOCA の入力画面（左）と母音パネルによる抑揚の操作方法（右）

上させるためには、

- モーラ音声の抑揚変化を固定されたパターンに限定せず、連続的に自由に变化させることを可能にする
- 事前に収録した音声の登録機能を付与することで、出力音声の個性を変更できるようにする

といった機能の実装が望まれる。

3. 改良型 VOCA の提案

従来 VOCA の表現力改善を図った VOCA の提案を行う。提案 VOCA の入力画面を図 2 の左に示す。子音パネルと母音パネルを順に押す操作は従来 VOCA と同じである。母音パネルによる抑揚の操作方法が異なり、左右にパネルを動かすと、モーラ音声のピッチの高低が連続的に、リアルタイムで変化する（図 2 右）。また、モーラ間のピッチの高さを滑らかに接続するため、母音パネルの左右の位置は全母音で連動して動くようになっている。（あるモーラ音声の出力を終え、パネルを離しても左右の位置が維持さ

れるため、次のモーラ音声の高さが、前のモーラ音声の高さと同じになる。)

また、ユーザー自身の声など、希望する話者の音声の録音・登録機能を有している。システムに「あ行」「か行」などの行ごとに発声した音声を録音すると、音声認識を利用した強制切り出しによって、モーラごとの音声に分割される。登録されるモーラは 110 種である。それぞれのモーラ音声に対して分析を行い、音声合成用のパラメータ（スペクトル包絡情報、非周期性指標）を求め、データベースに登録する。音声出力を行う際には、登録されたパラメータで合成フィルタを構築し、そこに母音パネルで指定した基本周波数の音源を入力することで、出力音声を得られる。母音はボタンの押下時間に応じて長さが変化するため、長く音声を出力する場合には、母音の中間部分の合成パラメータを繰り返し使用する。

音声の分析とリアルタイム合成は、音声分析合成システム WORLD [2] を利用して実装を行った。

4. 提案 VOCA の有効性の評価

提案 VOCA で新たに導入した、抑揚操作と個人性反映のそれぞれの有用性を検証するため、被験者による評価実験を行った。被験者は大学生 5 名であり、実際に自分が発声障がいを持ったことを想定して、評価を行ってもらった。

4.1 抑揚操作に関する評価

被験者に従来 VOCA と提案 VOCA の両方を使用した対話を体験してもらい、それぞれのシステムの有効性を評価してもらった。対話は、「事前にシナリオを用意した一問一答形式の模擬対話」として、従来 VOCA と提案 VOCA の使用時に発話量や発声内容が大きく変わらないように制御した。また、両 VOCA の操作に慣れてもらうため、対話を行う前にそれぞれ 40 分程度の練習時間を設けている。

評価項目は、「話し方の幅広さ」「感情・態度の表現力」「出力音声の自然性」「操作の容易さ」「対話の円滑さ」「総合評価」であり、それぞれについて最低を 1、最高を 5 とする 5 段階で評価してもらった。表 1 にそれぞれの項目に対する平均評価値の比較を示す。このうち、危険率 5% で有意差が確認されたものに「*」を示す。これをみると、操作性を維持しながら、「話し方の幅広さ」「感情・態度の表現力」といった点の表現の改善があることが確認された。

表 1 韻律操作に関する評価結果

評価項目	従来 VOCA	提案 VOCA
話し方の幅広さ	3.4	4.6*
感情・態度の表現力	3.2	3.8*
出力音声の自然性	2.4	3.6
操作の容易さ	3.8	4.0
対話の円滑さ	3.2	4.0*
総合評価	3.0	4.4*

表 2 個人性反映に関する評価結果

評価項目	平均評価値
個人性の再現度	3.8
音声の明瞭性	3.5

4.2 個人性反映に関する評価

個人性反映の評価を行うため、提案 VOCA に被験者自信の声を登録した場合の「個人性の再現度」「音声の明瞭性」について 5 段階評価（最低 1、最高 5）を行ってもらった。被験者は男子大学生 10 名である。評価に際しては、模擬対話課題を実施してもらっている。表 2 に両項目の平均評価値を示す。明瞭性に関してはやや低い評価となっているが、再現度は 4 に近い評価となっており、その有効性を確認することができる。

5. まとめ

本研究では、我々が提案・開発を行っているモーラ単位で音声出力を行う形式の VOCA について、従来のものから表現力向上を図った改良版の VOCA の提案を行った。具体的には「連続的な抑揚変化」と「個人性の反映」を、リアルタイム音声分析合成を利用して実装している。評価実験の結果、提案 VOCA の有効性を確認することができた。

今後の課題としては、実際の発声障がい者の方に被験者になってもらった上での提案 VOCA の評価などが挙げられる。

参考文献

- [1] 仲戸川, 岩野, “モーラ単位音声出力を用いた発声障がい者のための会話支援システム,” 日本音響学会秋季研究発表会講演論文集, 1-2-4, pp. 1385-1386, 2013.
- [2] M. Morise, et al., “WORLD: a vocoder-based high-quality speech synthesis system for real-time applications,” IEICE transactions on information and systems, vol. E99-D, no. 7, pp. 1877-1884, 2016.